

The MMI Ontology Registry and Repository: A Portal for Marine Metadata Interoperability

Carlos Rueda – Luis Bermudez – Janet Fredericks

Monterey Bay Aquarium Research Institute
Southeastern Universities Research Association
Woods Hole Oceanographic Institution



Oceans'09 MTS/IEEE Biloxi – October 27, 2009

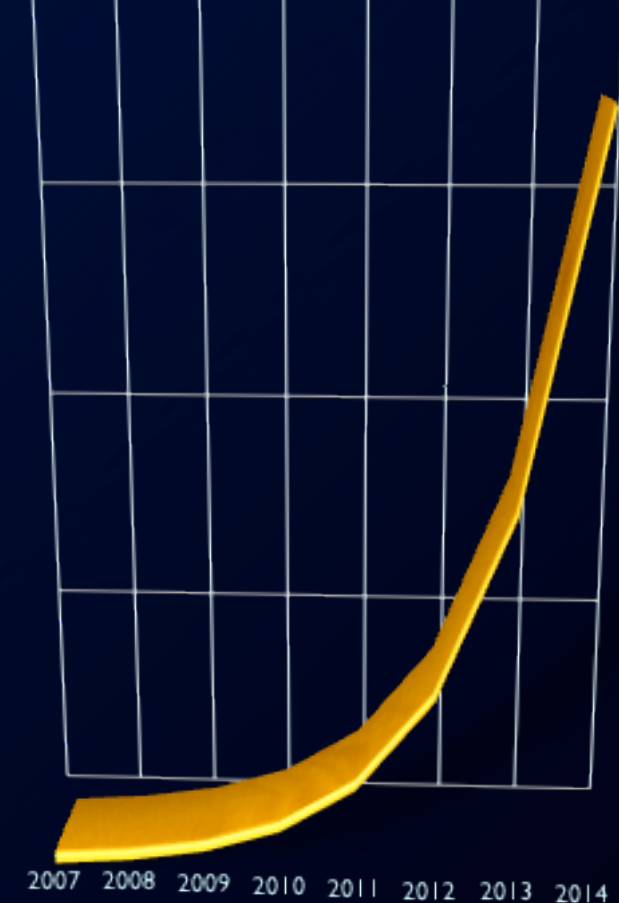
Outline

- The Challenge and the Approach
- MMI Ontology Registry and Repository
- Enhancing Standards with Semantics
- Ongoing and Future Work

Got data?

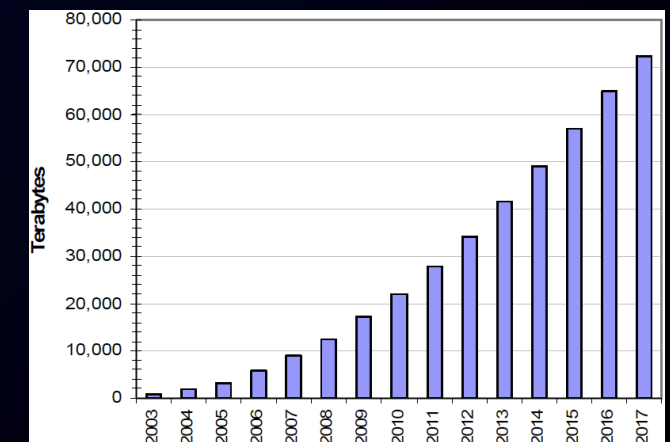
- Doubling each year in the physical, Earth, and life sciences*

- Sloan Digital Sky Survey, GDB Human Genome Database, Global Biodiversity Information Facility, ...



- Earth sciences

- Dozens of terabytes per day (NASA, NOAA)
- 2002–2017 NOAA's data holdings expected to grow by a factor of 100 to 74 petabytes**



* 2020 Computing: Science in an exponential world. Szalay and Gray. Nature (2006)

** NOAA The Nation's Environmental Data: Treasures at Risk (2001)

But, are we exploiting the data as well?

- Sharing data effectively?
- Discovering data easily?
- Easily assimilating and integrating data?
- Taking action to facilitate all of the above?

Interoperability

- Standards efforts
 - International Organization for Standardization (ISO)
 - Open Geospatial Consortium (OGC)
 - World Wide Web Consortium (W3C)
- But there still remains a gap for the effective exploitation of data: *semantic heterogeneity*

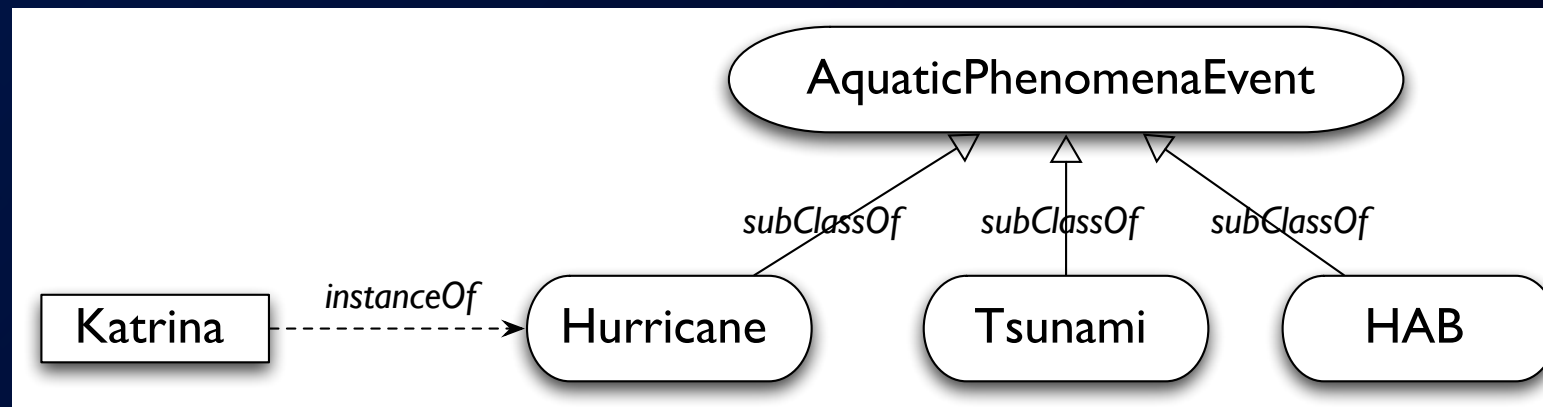
Semantic Web

Linked data*

- Make the content on the web interconnected in a meaningful way for both humans and **machines**
- Key technologies:
 - RDF and Ontologies
 - Formalized resource descriptions and knowledge representation
 - URI
 - Uniform identification of resources
 - HTTP
 - Hypertext Transfer Protocol

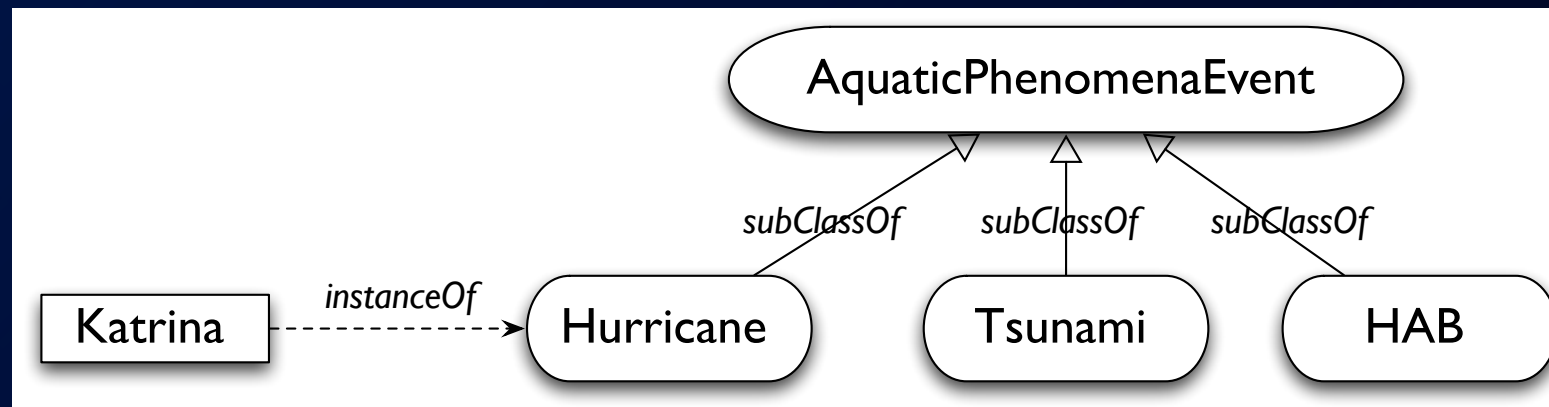
* Tim Berners-Lee. “Linked Data.” <http://www.w3.org/DesignIssues/LinkedData.html> (2006)

Ontologies



- “Hurricanes,” “Tsunamis,” and “Harmful algal blooms” are “aquatic phenomena events”
- “Katrina” is a particular “Hurricane” event

Ontologies



- “Hurricanes,” “Tsunamis,” and “Harmful algal blooms” are “aquatic phenomena events”
- “Katrina” is a particular “Hurricane” event
- However, different communities may use different terminologies
 - Eg., use the term “red tide” instead of “harmful algal bloom” to designate the same phenomenon
- Need mechanisms to link terms across multiple vocabularies with a variety of possible relationships

MMI Ontology Registry and Repository

MMI Ontology Registry and Repository

- Registry
 - MMI ORR is a catalog of (pointers to) ontologies and associated metadata

MMI Ontology Registry and Repository

- Registry
 - MMI ORR is a catalog of (pointers to) ontologies and associated metadata
- Repository
 - MMI ORR contains the registered ontologies

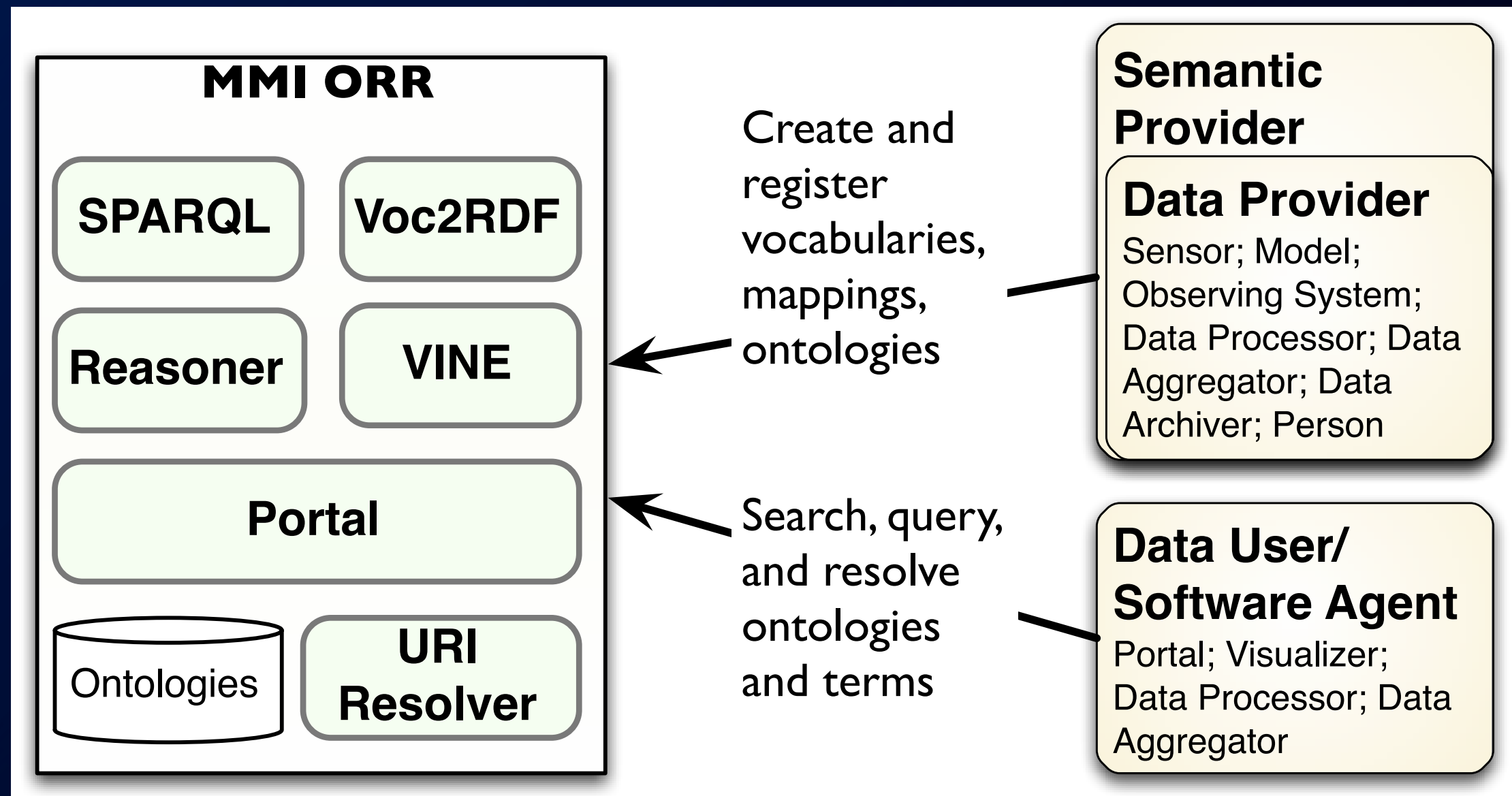
MMI Ontology Registry and Repository

- Registry
 - MMI ORR is a catalog of (pointers to) ontologies and associated metadata
- Repository
 - MMI ORR contains the registered ontologies

MMI Ontology Registry and Repository

- Registry
 - MMI ORR is a catalog of (pointers to) ontologies and associated metadata
- Repository
 - MMI ORR contains the registered ontologies
- But wait, there's more ...

MMI Ontology Registry and Repository




MMI Ontology Registry and Repository

Some key requirements

- Controlled vocabulary creation
 - Registration/storage of vocabularies
 - Metadata association
 - Versioning
 - Semantic query support
 - Term mapping
 - Mapping repository
 - Vocabulary and term URI resolution
- **Easy-to-use** tools for creation of vocabularies and mappings
- **Semantic mediation services** to diverse data portals and tools

MMI ORR Portal



Marine Metadata Interoperability
Ontology Registry and Repository
alpha

[caruedagm](#) | [Sign out](#) | [Help](#)
[Search terms](#)

☒ All ontologies
Registered by me
☒ Registered by:
☒ Type
mapping
vocabulary
☒ Authority
agu
argo
bodc
cdip
cencoos
cf
codesmelltest
cuahsi

URI	Name	Author	Version
▶ http://mmisw.org/ont/agu/indexterm	AGU Index Terms	AGU	20090604T210443
▶ http://mmisw.org/ont/CUAHSI/AquaBiologicalCompound	Aqua Biological Compound	mattr	20090914T213519
▶ http://mmisw.org/ont/CUAHSI/AquaBioCore	Aqua Biological Core	mattr	20090914T213352
▶ http://mmisw.org/ont/argo/instrument	ARGO Instruments	ARGO	20090605T194103
▶ http://mmisw.org/ont/argo/parameter	ARGO Parameters	ARGO	20090605T194513
▶ http://mmisw.org/ont/argo/qualityFlag	Argo QA/QC Flags	Stephanie Watson	20081116T040146
▶ http://mmisw.org/ont/CUAHSI/AtmosphericHydrologicCompound	Atmospheric Hydrologic Compound	mattr	20090914T213157
▶ http://mmisw.org/ont/CUAHSI/AtmosphericHydrologicCore	Atmospheric Hydrologic Core	mattr	20090914T212958
▶ http://mmisw.org/ont/CUAHSI/AtmosphericHydrologicSynthesis	Atmospheric Hydrologic Synthesis	mattr	20090914T212842
▶ http://mmisw.org/ont/mmi/authority	Authority Vocabulary	MMI	20090317T183911

Vocabulary Editor (Voc2RDF)

Metadata

Contents

General Usage/License/Permissions Original source

These fields capture general information about this ontology, who created it, and where it came from. The two attributes "Resource type" and "Authority abbreviation" are used to construct the URIs for the vocabulary and terms. In our system, they can contain only letters, numbers, underscores, and (not recommended) hyphens, and begin with a letter.

Example Reset

Resource type:*

Choose

URI of resource type:

Full title:*

BODC Parameter- Group Dictionary

Acronym:

parametergroupcode

Content creator:*

Ontology creator:*

BODC

Brief description:*

Parameter Group codes
More information: http://wwwtest.bodc.ac.uk/data/codes_and_formats/parameter_codes/

Keywords:

Link to original vocabulary:

http://www.bodc.ac.uk/data/codes_and_formats/parameter_codes/

Link to documentation:

Authority abbreviation:*

bodc

Choose

Class name: Group				
	id	short_name	definition	long_name
1	D000	D000	Determinations of the abundance of diatom remains in sediment linked to taxonomic identifications that may be mapped to entities in the ITIS taxonomy. These may be at any taxonomic level from sub-species upwards	Diatom taxonomy-related abundance per unit mass of sediment
2	OHWC	WC_OthHcar	'Other' (not alkanes, alkenes, alkynes or PAHs) hydrocarbon concentration parameters (including saturation of gaseous species) in all phases of the water column. Does not include parameters expressed per unit weight of SPM.	Concentration of 'other' hydrocarbons in the water column
3	ZCTC	Zoo_taxa_C_biomass	Zooplankton (mesozooplankton plus larger pelagic animals excluding fish, reptiles and mammals) carbon biomass parameters presented at the level of taxa that may be mapped to entities in the ITIS taxonomy. These may be at any taxonomic level from sub-species upwards.	Zooplankton taxonomy-related biomass expressed as carbon per unit volume of the water column
4	ASLV	SeaLvl	Measurements of the displacement of the water column surface from a fixed, stable reference	Sea level
5	PCHW	WC_PAH	PAH concentrations in all phases of the water column. This group includes concentrations per unit volume of the water column in the particulate phase, but not concentrations per unit weight of SPM.	Concentration of polycyclic aromatic hydrocarbons (PAHs) in the water column
6	OPBS	OptBS		Optical backscatter
7	UREA	Urea	Urea concentration parameters (including statistical parameters such as standard deviation) in the water column	Urea concentration parameters in the water column

Term Mapping Editor (VINE)

Working ontologies: Add...

A: <http://mmisw.org/ont/q2o/qualityFlag> -- Q2O Quality Control Flags

B: <http://mmisw.org/ont/igoss/qualityFlag> -- Integrated Global Ocean Services System (IGOSS) QA/QC Flag

Search the following ontologies: A B

Search for: [] REGEX

Select: All None Selected: 1 out of 8 element(s)

☒ A:/fail

<http://mmisw.org/ont/q2o/qualityFlag/fail>

label: fail

id: fail

code: 0

code_type: boolean

definition: test failed

relationship: <http://mmisw.org/ont/q2o/qcCategory/fail>

type: QualityFlag

☐ A:/id

Search the following ontologies: A B

Search for: [] REGEX

Select: All None Selected: 1 out of 14 element(s)

☐ B:/_6

☒ B:/_3

http://mmisw.org/ont/igoss/qualityFlag/_3

label: 3

code: 3

description: The element is probably bad

source_notes: Integrated Global Ocean Services System from IODE Summary Spreadsheet QualityFlags.xls

type: Qualityflag

☐ B:/_1

Config

Select: All None

Mappings:

A:/fail. = B:/_3.

Legend:


- = Exact match
- ≈ Close match
- < Narrower than
- > Broader than
- ~ Related match

Search

- SPARQL: Query Language for RDF
- Output formats: RDF/XML, N3, JSON, CSV

Search

- SPARQL: Query Language for RDF
- Output formats: RDF/XML, N3, JSON, CSV

Search terms containing: <input type="text" value="adcp"/>  Use OR to separate alternative keywords		
subject	predicate	object
http://mmisw.org/ont/bodc/instrument/Acoustic_doppler_current_profiler_ADCP	http://mmisw.org/ont/bodc/instrument/name	Acoustic doppler current profiler (ADCP)
http://mmisw.org/ont/bodc/instrument/Acoustic_doppler_current_profiler_ADCP	http://www.w3.org/2000/01/rdf-schema#label	Acoustic doppler current profiler (ADCP)
http://mmisw.org/ont/bodc/instrument/Lowered_ADCP	http://mmisw.org/ont/bodc/instrument/name	Lowered ADCP
http://mmisw.org/ont/bodc/instrument/Lowered_ADCP	http://www.w3.org/2000/01/rdf-schema#label	Lowered ADCP
http://mmisw.org/ont/bodc/parametergroupcode/LERR	http://mmisw.org/ont/bodc/parametergroupcode/definition	Parameters as logged by measuring devices, including raw counts/voltages and ADCP relative velocities
http://mmisw.org/ont/ecs/device/SONAR_ADCP	http://mmisw.org/ont/ecs/device/subtype	ADCP

Search

- SPARQL: Query Language for RDF
- Output formats: RDF/XML, N3, JSON, CSV

Search terms

http://mmisw.org/ont/bodc/instrumentcode/Lowered_ADCP

Submit

SPARQL query:

```

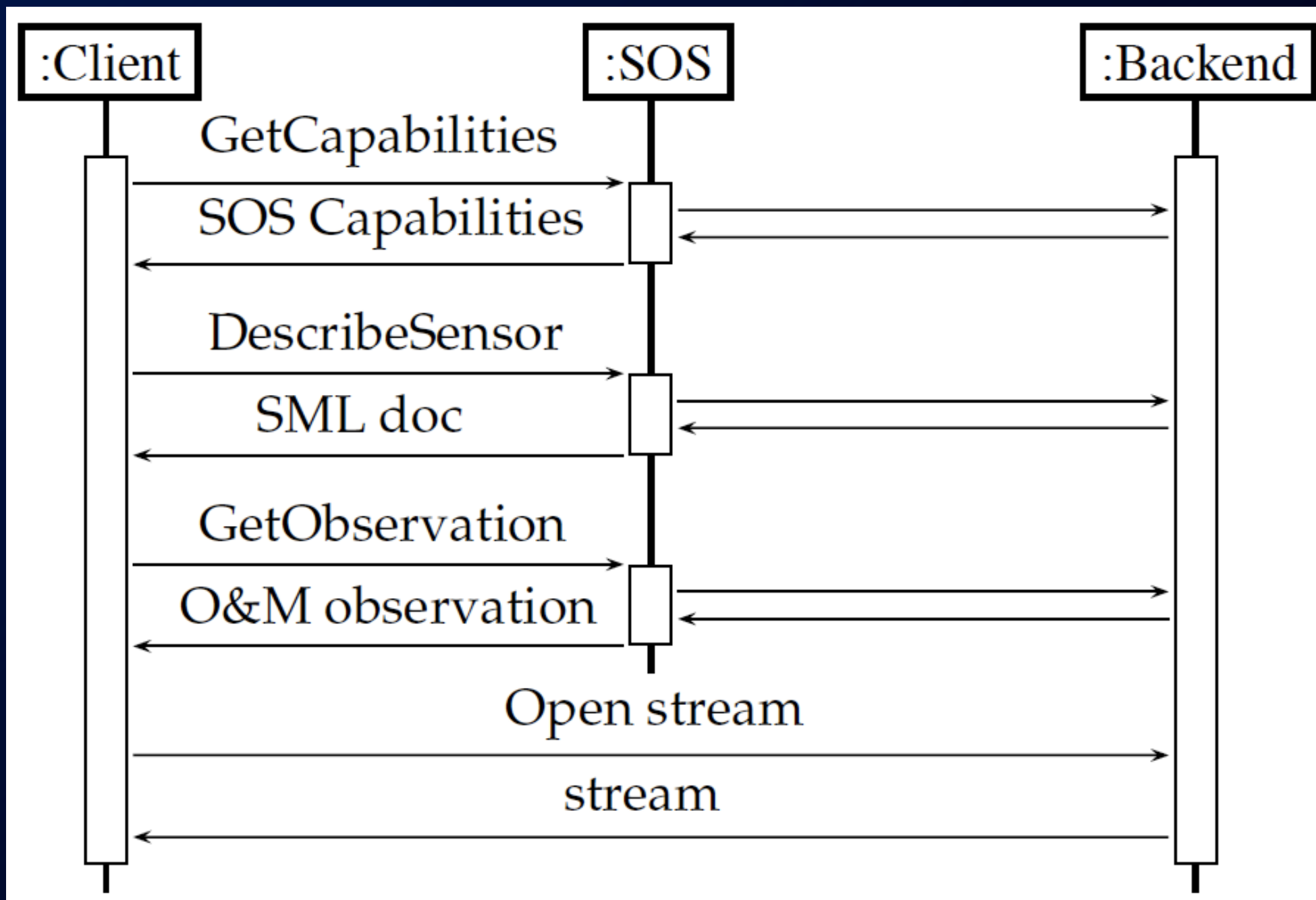
PREFIX dc: <http://purl.org/dc/elements/1.1/>
SELECT ?resource ?creator
WHERE { ?resource dc:creator ?creator. } LIMIT 12
        
```

resource	creator
http://mmisw.org/ont/igoss/20081113T203318/qualityFlag	Stephanie Watson
http://mmisw.org/ont/gcoos/20081118T044314/sensor	YIXIN LUO
http://mmisw.org/ont/mmi/20081119T135714/usjgofs_hydro	Cyndy Chandler
http://mmisw.org/ont/q2o/20081117T133042/qualityFlag	Sara Haines
http://mmisw.org/ont/qartod/20081113T212733/qualityFlag	Stephanie Watson
http://mmisw.org/ont/cf/20081116T050933/parameter	Luis Bermudez
http://mmisw.org/ont/ecs/20081120T164852/device_map	Amilan
http://mmisw.org/ont/cencoos/20081116T071031/parameter	Dale Robinson
http://mmisw.org/ont/ndbc/20081114T205115/qualityFlag_hard	Stephanie Watson
http://mmisw.org/ont/seadatanet/20081116T060423/qualityFlag	Stephanie Watson
http://mmisw.org/ont/mmittest/20081117T202337/qualityFlag	Carlos Rueda
http://mmisw.org/ont/cencoos/20081119T231517/parameter	Dale Robinson

Open Geospatial Consortium Sensor Web Enablement (OGC SWE)

- “Sensor Web”
discover, access, command sensors of all types
- SensorML
 - to describe sensors and sensor platforms
- O&M
 - to encode observations and measurements
- SOS: Sensor Observation Service

Sensor Observation Service

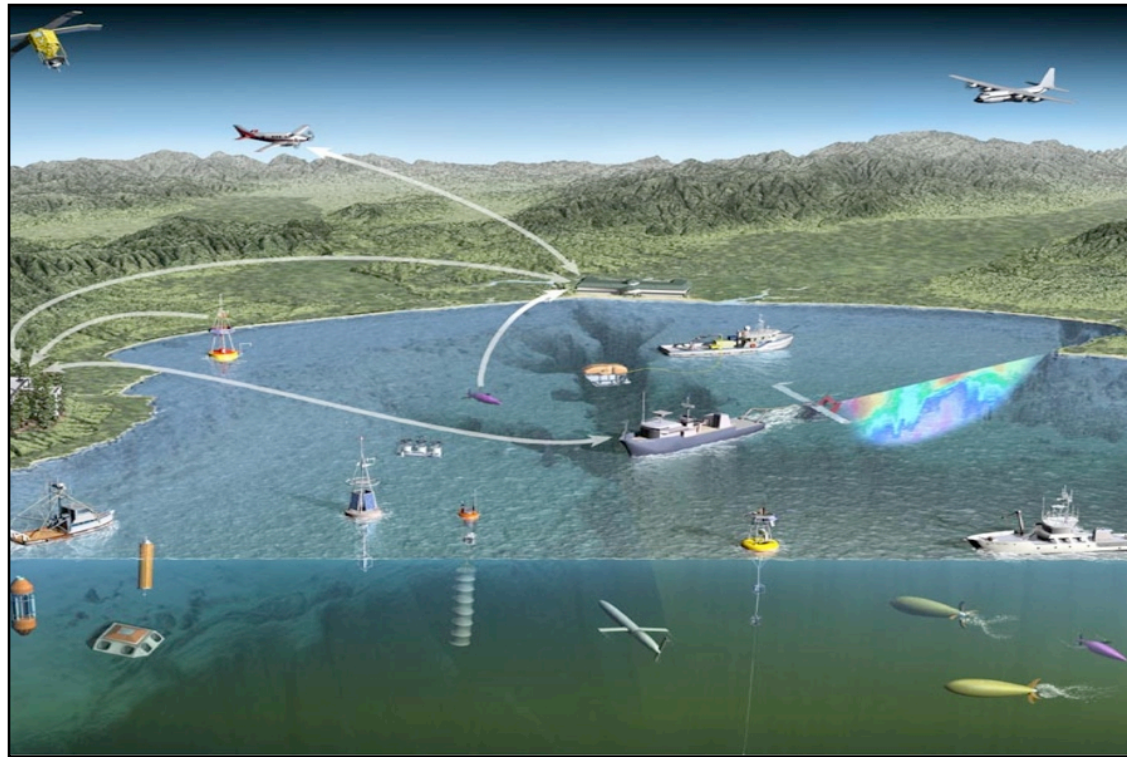


Enhancing SOS with Semantics Linked (meta)data

- Example with a SOS observation offering

```
- <sos:ObservationOffering gml:id="observationOffering_1455">
  <gml:description/>
  - <gml:boundedBy>
    - <gml:Envelope srsName="urn:ogc:def:crs:EPSG:6.5:4326">
      <gml:lowerCorner>36.69623 -122.39965</gml:lowerCorner>
      <gml:upperCorner>36.69623 -122.39965</gml:upperCorner>
    </gml:Envelope>
  </gml:boundedBy>
  - <sos:time>
    - <gml:TimePeriod gml:id="timePeriod3">
      <gml:beginPosition>2008-06-09T09:36:19Z</gml:beginPosition>
      <gml:endPosition>2008-06-10T02:06:21Z</gml:endPosition>
    </gml:TimePeriod>
  </sos:time>
  <sos:procedure xlink:href="urn:mbari:org:device:1455"/>
  <sos:observedProperty xlink:href="http://mmisw.org/ont/cf/parameter/sea_water_temperature"/>
  <sos:observedProperty xlink:href="http://mmisw.org/ont/cf/parameter/conductivity"/>
  <sos:observedProperty xlink:href="http://mmisw.org/ont/cf/parameter/pressure"/>
  <sos:observedProperty xlink:href="http://mmisw.org/ont/cf/parameter/sea_water_salinity"/>
  <sos:featureOfInterest xlink:href="http://mmisw.org/mmi/20080516/system#EarthRealm"/>
  <sos:responseFormat>text/xml; subtype="om/1.0.0"</sos:responseFormat>
  <sos:responseMode>inline</sos:responseMode>
</sos:ObservationOffering>
```


OGC Ocean Science Interoperability Experiment (OSIE)



Initiative to advance standards
for interoperability of ocean
observing systems

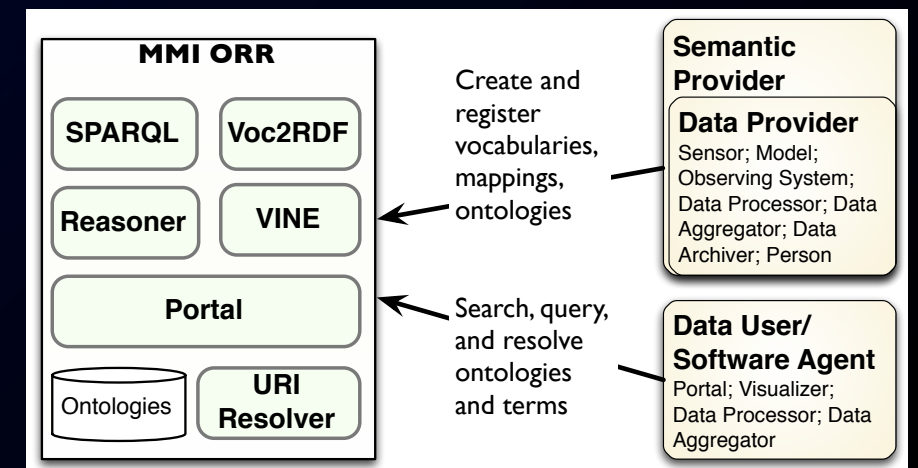


OGC Ocean Science Interoperability Experiment (OSIE)

- Use case: Find sensors/systems/data given:
 - sensor/measurement types, region of interest, time period, ...
 - free text or keywords
 - Include results for related terms (broader, narrower)
- **Categorization of sensors and observations**
 - supports the discovery functionality
 - based on ontologies

The Overall Process

- Portal
 - Creates/uses ontology representing the portal categories
- Data providers
 - Create/uses ontologies for the concepts used in their services
 - Create/uses ontology with mappings between service provider's terms and portal categories
- Portal and clients in general
 - Query the MMI ORR



Conclusions

- MMI ORR allows data providers and users to include, use, and exploit semantic information in real world applications
- OGC Sensor Web Enablement services can be enriched with semantic references that are resolvable against the MMI ORR
- Marine science data interoperability can be realized with semantic web technologies

Ongoing and future work

- **End-to-end semantic solutions for the Marine and Earth science communities**
 - Semantically enabled tools
 - Device and observations ontologies
- **MMI ORR**
 - Collaborative features
 - Enhance mappings support
 - Integration with external ontology tools

- <http://mmisw.org/orr> – MMI ORR
- <http://marinemetadata.org> – Marine Metadata Interoperability Project

Thank you!

Carlos Rueda – carueda@mbari.org
Luis Bermudez – bermudez@sura.org
Janet Fredericks – jfredericks@whoi.edu



Registration Options

- Fully hosted
 - mmisw.org-based URIs
 - Direct resolution of URIs
 - Versioning
- Re-hosted
 - Original namespace preserved
 - Indirect resolution of URIs
- Indexed
 - Ontology just incorporated in knowledge base

`http://mmisw.org/ont/mmi/20081203/parameter/sst`

Root

Authority

Version

Vocabulary

Term